2008

# Sound and immersion in the first-person shooter: mixed measurement of the player's sonic experience.

Mark Grimshaw
*University of Bolton*, m.n.grimshaw@bolton.ac.uk

Craig A. Lindley

Lennart Nacke

# Sound and Immersion in the First-Person Shooter: Mixed Measurement of the Player's Sonic Experience

Mark Grimshaw (mark@wikindx.com)
Craig A. Lindley (craig.lindley@bth.se)
Lennart Nacke (Lennart.Nacke@bth.se)

---

**Abstract.** Player immersion is the holy grail of computer game designers particularly in environments such as those found in first-person shooters. However, little is understood about the processes of immersion and much is assumed. This is certainly the case with sound and its immersive potential. Some theoretical work explores this sonic relationship but little experimental data exist to either confirm or invalidate existing theories and assumptions.

This paper summarizes and reports on the results of a preliminary psychophysiological experiment to measure human arousal and valence in the context of sound and immersion in first-person shooter computer games. It is conducted in the context of a larger set of psychophysiological investigations assessing the nature of the player experience and is the first in a series of systematic experiments investigating the player's relationship to sound in the genre. In addition to answering questionnaires, participants were required to play a bespoke *Half-Life 2* level whilst being measured with electroencephalography, electrocardiography, electromyography, galvanic skin response and eye tracking equipment. We hypothesize that subjective responses correlated with objective measurements provide a more accurate assessment of the player's physical arousal and emotional valence and that changes in these factors may be mapped to subjective states of immersion in first-person shooter computer games.

---

## Introduction

An increasing amount of research in games studies and games technology deals with presence in virtual environments and, of more interest for the purposes of our research, player immersion in digital game worlds. Player immersion may be said to be the holy grail of digital game design particularly in the types of game environment found in the First-Person Shooter (FPS) genre. This type of game is typically exemplified by the run 'n' gun sub-genre (games such as the *Doom*, [1] *Quake* [2] and *Half-Life* [3] series – even though the latter has an unusually strong focus on narrative) which is visually characterized by a hand or pair of hands holding a weapon on screen and conceptually characterized as 'the hunter and the hunted'. The intention is that the player identifies with the game character whose hands, the player's virtual prostheses, are seen receding into the game environment. [4] This identification with the character, and the use of hands only, provides a first-person perspective with which it is proposed that, visually, player immersion in the game world derives from the player 'becoming' the game character, in the sense of the player having the experience of acting 'within' the game world. This sense of immersion is strengthened further through the player's actions having a non-trivial effect on the environment and game play. [5] For example, operating the game interface (the computer mouse or keyboard, for instance) may cause the image on the screen to change: the weapon may recoil and flash or an on-screen animation indicates the weapon reloading. The player perceives movement through the 3-dimensional world of the game because visual artifacts rotate, magnify and diminish or appear and disappear on the screen.

There are a variety of definitions of immersion in computer games. Kearney and Pivec claim that immersion provides the motivation or flow required for the player to be repeatedly engaged with the game [6] while Ermi and Mäyrä, paraphrasing Pine and Gilmore, [7] state that "immersion means becoming physically or virtually a part of the experience itself". They also distinguish between different forms of immersion: sensory, imaginative and challenge-based immersion. [8] Murray suggests that immersion is a participatory activity [9] and McMahan provides three conditions for immersion: "[T]he user's expectations of the game or environment must match the environment's conventions fairly closely […] the user's actions must have a non-trivial impact on the environment [and] the conventions of the world must be consistent". [5] For McMahan, two factors influencing immersion are the level of social realism and the level of perceptual realism. Garcia claims that "[i]n the most immersing environments reminders of the structural level of the game are gone" [10] while Carr provides two categories of immersion: perceptual (when the participant's senses are monopolized by the experience) and psychological (an imaginative or mental absorption through which the participant becomes engrossed in the experience). [11]

In the context of this paper, other work mentions or focuses upon the role of sound in facilitating player immersion in the game world. Laurel makes the case that the "[t]ight linkage between visual, kinaesthetic, and auditory modalities" is key to the sense of immersion. [12] Jørgensen believes that players are immersed in an auditory world through the use of realistic audio samples [13] while Murphy and Pitt make similar claims for spatial sound. [14] Autopoiesis and acoustic ecologies have been used to model player immersion through sound in FPS games [4] and Grimshaw and Schott provide a range of conceptual tools with which to analyze the immersive functions of game sound. [15] Some authors make a distinction between modes of immersion, particularly where immersion is enabled through the spatial qualities of the sound: Stockburger implies that the player is physically immersed in the game sound [16] and this is amplified and explicitly stated by Grimshaw. [17] This physical sonic immersion has also been observed for film audiences and the concept transferred to the design of sound for FPS games and simulators. [18]

Most of the work cited above is theoretical and, where authors describe the immersive potential of sound in computer games, there is the assumption that sounds, more sounds, realistic sounds, spatial sounds all inexorably and incontrovertibly equate to greater player immersion. This may well be the case but the assumption lacks thorough evidence to support the various concepts of immersion outlined above. Attempts to provide evidence include Jørgensen, who uses player surveys, [13] and Shilling, Zyda and Wardynski. [18] The latter paper is of particular interest to this work because it not only explores the use of sound in an FPS game/simulation (*America's Army* [19]) but it also attempts to objectively measure the player's emotional arousal through the use of temperature, electrodermal response and heart-rate measurements. However, although the authors state that "emotional arousal has a positive impact on [the] sense of immersion in virtual environments" and that the precise conjunction of a sound and an action seen on the screen is "crucial for immersing the player", the paper is a description of their attempts to introduce, and amplify, *emotion* within the game environment through sound rather than an attempt to effect and measure *immersion*. The link between emotional arousal and immersion is assumed and so the relationship between sound and player immersion remains undefined in objective terms.

Emotions are a central part of the game experience, motivating the conscious cognitive judgments and decisions made during gameplay. Psychophysiological investigations suggest that at least some emotional states may be quantitatively characterized via physiological measurements. Specific types of measurement of different physiological responses (such as GSR, EMG, ECG and EEG, as described below) are not by themselves reliable indicators of well-characterized feelings; [20, 21] a *de rigueur* cross-correlation of all measurements is crucial to identify the emotional meaning of different patterns in the responses. Moreover, the often described many-to-one relation between psychological processing and physiological response [20] allows for psychophysiological measures to be linked to a number of psychological structures (for example, attention, emotion, information processing). Using a response profile for a set of physiological variables enables researchers to go into more detail with their analysis and allows a better correlation of response profile and psychological event. [21] The crucial issue here is the correlation of patterns of measurement characteristics for a set of different measures with subjective characterizations of experience such as emotion and feelings (for example, the feeling of immersion in gameplay).

Facial electromyography (EMG) is a direct measure of electrical activity involved in facial muscle contractions; EMG provides information on emotional expression via facial muscle activation (even though a facial expression may not be visually observable) and can be considered as a useful external measure for hedonic valence (that is, degree of pleasure/displeasure). [22] Positive emotions are indexed by high activity at the zygomaticus major (cheek muscle) and orbicularis oculi (periocular muscle) regions. In contrast to this, negative emotions are associated with high activity at the corrugator supercilii (brow muscle) regions. This makes facial EMG suitable for mapping emotions to the valence dimension in the two-dimensional space described in Lang's dimensional theory of emotion. [22] This dimension reflects the degree of pleasantness of an affective experience. The other dimension, the arousal dimension, depicts the activation level linked to an emotionally affective experience ranging from calmness to extreme excitement. In this kind of dimensional theory of emotion, emotional categories found in everyday language (for example, happiness, joy, depression, anger) are interpreted as correlating with different ratios of valence and arousal, hence being mappable within a space defined by orthogonal axes representing degrees of valence and arousal, respectively. For example, depression may be represented by low valence and low arousal, while joy may be represented by high valence and high arousal.

Arousal is commonly measured using galvanic skin response (GSR), also known as skin conductance. [23] The conductance of the skin is directly related to the production of sweat in the eccrine sweat glands, which is entirely controlled by the human sympathetic nervous system. Increased sweat gland activity is thus directly related to electrical skin conductance. Hence, measuring both GSR and EMG provides sufficient data to provide an interpretation of the emotional state of a game player in real time, according to a phasic emotional model.

This paper describes and analyzes the results of a preliminary experiment that investigates the role of sound in enabling player immersion in the FPS game *Half-Life 2*. [24] The investigation is designed to provide both subjective responses (through the use of questionnaires) and objective measurements (through the use of electromyography (EMG), galvanic skin response (GSR), electroencephalography (EEG), electrocardiography (ECG) and eye tracking equipment). The overall aim of the experiment is to find external (that is, objective) measures that may be reliably correlated with subjective experiences assessed via questionnaires in order to provide more detailed descriptions of the emotional experience of game players during gameplay, both in the degree of emotions experienced and in the timescale of emotional changes and modulations. It is further hoped that this method may lead to real-time measures for states of immersion of players playing first-person shooter computer games. Finally, correlating discriminations within psychophysiological data with different categories of immersion can provide at least one method for validating those categorizations. The experiment is preliminary since the psychophysiological characterization of states of immersion is not yet well developed.

The study further aims to provide a psychophysiologically-based answer to the assumption that sound plays a role in enabling the immersion of the player in the FPS game world. If the results of the experiment provide a positive answer, that sound does indeed play this role, it is envisaged that future experiments, using a similar methodology, will be designed to investigate more specific questions about the relationship between the player and sound in FPS games.

The experiment was conducted in May 2008 in the Game and Media Arts Laboratory at Blekinge Institute of Technology (BTH) in Sweden. The investigation of sound formed part of a larger psychophysiological investigation into the nature of the player experience in computer games. This paper is also limited to the analysis of GSR, EMG and questionnaire data. Further analysis taking into account the other data types is ongoing.

## Method

Subjects played a *Half-Life 2* game mod especially designed for a short immersive playing time of maximum 10 minutes. The game mod was played four times with different sound modalities and physiological responses were measured together with questionnaires (assessing subjective responses) for each modality.

### 2.1 Design

The game sessions were played under four different conditions, corresponding to the permutations of the independent variable (sound modality): playing with diegetic game sounds (normal sounds), playing with speakers completely turned off (no sounds, no music), playing with diegetic game sounds and an additional music loop (sounds and music), and playing with diegetic game sounds turned off and hearing only the music loop (only music). Participants played under each condition in a shifting order to eliminate repeated-measures effects (using a Latin Squares design). Physiological responses (as indicators of valence and arousal) were recorded for each session as well as questionnaire answers. Questionnaire item order was randomized for each participant using the open-source software LimeSurvey. [25]

### 2.2 Participants

Data were recorded from 36 students and employees, recruited from the three BTH University campuses and their age ranged between 18 and 41 ($M$ = 24, $SD$ = 4.89). 19.4% of all participants were female. When asked how frequently they play digital games, 50% answered that they play games every day, 22.2% play weekly, 22.2% play occasionally and only 5.6% play rarely or never. However, it should be noted that 62.1% of all the males play on a daily basis and 20.7% play weekly. In contrast to that, most of the females enjoyed playing only on an occasional (57.1% of all females) or weekly (28.6%) basis.

Out of all participants, 47.2 % considered themselves casual gamers, 38.9% said that they belong to the hardcore gamer demographic and 13.9% could not identify themselves with any of those. Nevertheless, no female participant considered herself to be a hardcore gamer and 71.4% of all females said they were casual gamers. Male gamers were more evenly distributed among hardcore (48.3%) and casual (41.4%) gamers: the larger percentage of males considering themselves hardcore players.

91.7% of the participants were right-handed and 50% were wearing glasses or contact lenses. 94.4% believed they had full hearing capacity (5.6% stated explicitly that they lack full hearing capacity). 69.4% had a preference for playing with a music track on. 44.4% preferred playing with surround sound speakers, while 33.3% opted for playing with stereo headphones. 11.1% liked playing with stereo speakers and the final 11.1% preferred surround sound headphones. 33.3% played an instrument. 13.8% played the piano or keyboard and 8.3% played the guitar. 41.7% saw themselves as hobby musicians – some people worked with sound recording and programming but did not play instruments.

66.7% of participants were enrolled as University students. 16.7% already had a Bachelor’s degree and 13.9% had a Master’s degree. 61.1% of the participants had already played the digital game *Half-Life 2* before, 30.6% played it between 10 and 40 hours and 58.3% played it on a PC, leaving only one participant who played it on an Xbox 360.

To estimate preconceptions of sound immersion, participants were also asked how important they considered sounds, in general, for first-person shooters (FPS). The results were rated on a 5-point scale ranging from 1 (*not important*) to 5 (*very important*). 55.6% claimed that sound was *very important* and 36.1% said it to be *important*. The term “immersive”, which was also part of the questionnaire items assessing sound immersion, was explained to participants beforehand as “the feeling of being encapsulated inside the game world and not feeling in front of a monitor anymore”. This was so phrased for reasons of lay intelligibility and deemed to be a synthesis of previous definitions of game immersion noted above, particularly those of Ermi and Mäyrä, Garcia and Carr. This is suitable for investigating whether immersion in a very general sense may be distinguishable in psychophysiological measurement features; if so, ongoing experiments may address the psychophysiological detection of finer distinctions within the broad category of immersion.

### 2.3 Apparatus

*Facial EMG.* We recorded the activity from left orbicularis oculi, corrugator supercilii, and zygomaticus major muscle regions, as recommended by Fridlund and Cacioppo, [26] using BioSemi flat-type active electrodes (11mm width, 17mm length, 4.5mm height) electrodes with sintered Ag-AgCl (silver/silver chloride) electrode pellets having a contact area 4 mm in diameter. The electrodes were filled with low impedance highly conductive Signa electrode gel (Parker Laboratories, Inc.). The raw EMG signal was recorded with the ActiveTwo AD-box at a sample rate of 2 kHz and using ActiView acquisition software.

*Galvanic skin response (GSR).* The impedance of the skin was measured using two passive Ag-AgCl (silver/silver chloride) Nihon Kohden electrodes (1 microamp, 512 Hz). The electrode pellets were filled with TD-246 skin conductance electrode paste (Med. Assoc. Inc.) and attached to the thenar and hypothenar eminences of the participant’s left hand.

*Video recording.* A Sony DCR-SR72E video camera (handycam) PAL was put on a tripod and positioned approximately 50 cm behind and slightly over the right shoulder of the player for observation of player movement and in-game activity. In addition, the video recordings served as a validation tool when psychophysiological data were visually inspected for artifacts and recording errors.

*Game experience survey.* Different components of game experience were measured using the game experience questionnaire (GEQ). [27] As shown in a previous study by Nacke and Lindley, [28] the GEQ components can assess experiential constructs of immersion, tension, competence, flow, negative affect, positive affect and challenge with apparently good reliability.

*Sound immersion.* Subjective player experience of sound immersion was measured using our own additional questionnaire items rated on a 5-point scale ranging from 1 (for example, *not immersive*) to 5 (for example, *extremely immersive*) for sessions where sound was audible. Specific sound questions included the following:

- How important are sounds in general for you in FPS games?
- *Diegetic Sounds:*
  How immersive were the following?
  - Background sounds
  - Sounds of opponents
  - Sounds that you produced yourself (player-produced sounds)

  How important was the sound for you in the level you just played?
- *No Sound, No Music:*
  How much did it bother you to play without sound?
- *Nondiegetic Music Only:*
  Did you miss the sound effects in this level? (Yes/No)

Figure 1: The eye-tracking screen, EMG sensors and electroencephalography array

Other apparatus used, but not included in this analysis (the data will form the subject of a future paper), were a Biosemi 32-channel EEG system and a Tobii 1750 eye tracker (cf. Figure 1).

### 2.4 Procedure

We conducted all experiments on weekdays in the time from 10:00 a.m. to 6:00 p.m. with each experimental session lasting approximately two hours. The experiments were advertized especially to graduate and undergraduate students. All participants were invited to the newly established Game and Media Arts Laboratory at Blekinge Institute of Technology, Sweden. After a brief description of the experimental procedure, each participant filled in two forms. The first one was a compulsory "informed consent" form (with a request not to take part in the experiment when suffering from epileptic seizures or game addiction). The next one was an optional photographic release form. Each participant had to complete an initial demographic and psychographic assessment questionnaire prior to the experiment, which was immediately checked for completeness. Participants were then seated in a comfortable office chair, which was adjusted according to their individual height, electrodes were attached and the participant was asked to rest and focus on a black cross on a grey background on the monitor. During this resting period of 3-5 minutes, physiological baseline recordings were taken.

Next, participants were seated in front of a high-end gaming computer, which used a 5.1 surround sound system for playback (*Half-Life 2* sound quality settings on 'high'), and were encouraged to get acquainted with the game controls for two minutes (using a non-stimulus game level) if they did not indicate *a priori* FPS experience. The participants played the same *Half-Life 2* game level four times for 10 minutes (or until completed) in a counter-balanced order to eliminate repeated-measures effects. As this was a preliminary experiment designed to produce broad subjective answers and objective measurements from which future, more refined experiments can be designed, the following broad sound on/off modalities were chosen:

1. the level with all diegetic sounds and nondiegetic music audible
2. the level with just diegetic sounds audible
3. the level with just nondiegetic music audible
4. the level with no sound or music audible

After each modality, participants were asked to report their subjective experiences using questionnaires. After completion of all modalities, participants were thanked for their participation and paid a small participation fee before they were escorted out of the lab.

### 2.5 Reduction and Analysis of Data

Recorded psychophysiological data were visually inspected using BESA (MEGIS Software GmbH, Germany) and EMG data were also filtered using a Low Cutoff Filter (30 Hz, Type: forward, Slope: 6 dB/oct) and a High Cutoff Filter (400 Hz, Type: zero phase, Slope: 48dB/oct). If data remained noisy, they were excluded from further analysis. EMG data were rectified and exported together with GSR data at a sampling interval of 0.49 ms to SPSS for further analysis.

Mean values for physiological responses were calculated for epochs of complete session times (varying between five and 10 minutes). Psychophysiological data were corrected for errors using log and ln transformations. After histogram inspection, data was assumed to be close to a normal distribution (without elimination of single outliers). Means were calculated for items of each of the seven GEQ questionnaire components (immersion, tension, competence, flow, negative affect, positive affect, and challenge).

To test statistical significance of the results, one-way repeated-measures ANOVAs were conducted in SPSS using sound modality as the within-subject factor for each of the three EMG measures (orbicularis oculi, corrugator supercilii, zygomaticus major), the galvanic skin response (GSR) and all GEQ components.

## Results

For each sound modality, the participants were asked a few assessment questions. This included 5-point scale ratings for the immersiveness of sounds (*5 = most immersive*) after they played the level with diegetic game sounds only. The background sounds were rated just below the median value (*M = 2.97, SD = 1.03*) and opponent-produced sounds were rated higher (*M = 3.81, SD = 1.06*) than player-produced sounds (*M = 3.14, SD = 1.44*).[1] The presence of sounds in this level was rated very important (*M = 4.17, SD = 1.03*). After playing without any sounds or music, most participants also claimed that it bothered them a lot (*M = 4.06, SD = 1.19*). For the "music only" modality, it was noted as well that 75% of the participants missed the sound effects.

| GEQ component | Sound & Music | Diegetic Sound Only* | Nondiegetic Music Only | No Sound or Music |
|---|---|---|---|---|
| Immersion | 1.53 (1.08) | 1.51 (1.01) | 1.13 (0.95) | 0.85 (0.79) |
| Tension | 0.97 (1.12) | 1.41 (0.81) | 1.94 (1.13) | 1.57 (1.14) |
| Competence | 2.18 (1.21) | 2.24 (0.93) | 1.88 (1.10) | 1.57 (1.06) |
| Flow | 2.10 (1.28) | 2.37 (0.94) | 1.72 (1.32) | 1.50 (1.12) |
| Negative affect | 0.86 (0.86) | 1.06 (0.62) | 1.63 (1.02) | 1.71 (1.10) |
| Positive affect | 2.06 (1.11) | 2.16 (0.93) | 1.61 (0.93) | 1.49 (1.00) |
| Challenge | 2.18 (1.08) | 2.19 (0.69) | 1.96 (0.85) | 1.60 (0.81) |

Table 1: Means (and standard deviations) of the GEQ components for the four test sound modalities[2]

Table 1 shows a comparison of GEQ mean scores. A comparison of these values shows that regardless of GEQ components, people gave higher ratings (except for tension and negative affect) when sound was present. The presence of

[1] Potentially supporting the theory of challenge-based immersion. [8]

[2] N = 36, N of items = 2(*5), GEQ scale has values between 0 and 4 (median = 2).

diegetic sound (whether combined with music or not) also seems to be an enabling factor of the subjective experience of challenge and flow — flow especially seems be experienced more easily with diegetic sounds.

The complete absence of sound seems to negatively influence the subjective feeling of immersion to a significant degree as it is the lowest rated item in this modality. With missing auditory feedback, there is also a decrease in the feeling of competence among all participants. The combined presence of sound and music seems to also have a soothing effect on play as ratings for tension and negative affect are very low under this modality. It is also the modality that has the highest score for the immersion item. However, it should be noted that music also seems to be somewhat distracting from game flow since flow ratings are higher when music is omitted and only diegetic sounds are presented.

For GEQ components Immersion ($\chi^2(5) = 3.49$, $p > .50$), Competence ($\chi^2(5) = 10.28$, $p > .05$), Negative Affect ($\chi^2(5) = 5.36$, $p > .30$), and Flow ($\chi^2(5) = 10.12$, $p > .05$), Mauchly's test indicated that the assumption of sphericity had been met, but for the remaining items Tension ($\chi^2(5) = 11.98$, $p < .05$), Positive Affect ($\chi^2(5) = 11.56$, $p < .05$) and Challenge ($\chi^2(5) = 23$, $p < .05$) it was violated. Therefore, degrees of freedom were corrected for the latter three using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .80$, $\varepsilon = .84$, and $\varepsilon = .71$).

Statistical significance was achieved for all components (Immersion: $F(3, 105) = 8.20$, $p < .001$), Competence: $F(3, 105) = 4.49$, $p < .01$), Negative Affect: $F(3, 105) = 9.75$, $p < .001$), Flow: $F(3, 105) = 9.42$, $p < .001$), Tension: $F(2.39, 83.73) = 7.85$, $p < .001$), Positive Affect: $F(2.52, 88.21) = 6.18$, $p < .01$), and Challenge: $F(2.14, 74.78) = 5.17$, $p < .01$)). These results show that the subjective game experience measured with the GEQ was significantly affected by the different sound modalities.

Table 2 shows a comparison of the normalized physiological responses. Negatively valenced arousal would be indexed by increased GSR and corrugator supercilii activity (with decreased zygomaticus major and orbicularis oculi activity). [29] This is not the case for any of the accumulated measurements shown. The only notable decrease of orbicularis oculi and zygomaticus major activity is shown under the no sound condition. However, corrugator supercilii activity is also decreased and galvanic skin response is somewhat consistent across conditions.

| Physiological Response | Sound & Music | Diegetic Sound Only | Nondiegetic Music Only | No Sound or Music |
|---|---|---|---|---|
| Orbicularis oculi (ln[μV]) | 1.85 (0.37) | 1.85 (0.37) | 1.86 (0.42) | 1.79 (0.31) |
| Corrugator supercilii (ln[μV]) | 1.94 (0.25) | 1.90 (0.27) | 1.95 (0.33) | 1.89 (0.26) |
| Zygomaticus major (ln[μV]) | 1.98 (0.40) | 2.00 (0.38) | 2.00 (0.43) | 1.94 (0.35) |
| Galvanic skin response (log[μS]) | 0.72 (0.18) | 0.73 (0.17) | 0.70 (0.18) | 0.72 (0.17) |

Table 2: Means (and standard deviations) for the corrected physiological measurements (EMG and GSR) under the different modalities[3]

Accordingly, Mauchly's test indicated that the assumption of sphericity had been violated for orbicularis oculi EMG means ($\chi^2(5) = 25.16$, $p < .001$), corrugator supercilii EMG means ($\chi^2(5) = 57.65$, $p < .001$), zygomaticus major EMG means ($\chi^2(5)= 16.43$, $p = .006$) and GSR means ($\chi^2(5)= 52.41$, $p < .001$). Hence, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity for EMG means ($\varepsilon = .63$, $\varepsilon = .45$, and $\varepsilon = .76$) and GSR means ($\varepsilon = .47$).

Nevertheless, neither EMG responses (orbicularis oculi: $F(1.90, 53.21) = 0.86$, $p > .40$, corrugator supercilii: $F(1.36, 38.02) = 0.66$, $p > .40$, zygomaticus major: $F(2.27, 63.58) = 0.61$, $p > .40$), nor GSR [$F(1.40, 39.05) = 0.68$, $p > .40$] could achieve statistical significance for the repeated measures design. The results of the ANOVA show that tonic measurements of physiological response from an accumulated game session were not significantly affected by different sound modalities.

## Discussion and Future work

This paper has described and analyzed the results of a preliminary experiment to measure the effect of FPS sound and music on player valence and arousal and to detect any possible correlations between measurable valence and arousal features and self- reported subjective experience.

There are two important and related results. Firstly, the data gathered from the subjective questionnaires (see Table 1) shows a significant statistical difference between the four modalities over the GEQ components. This is particularly the case with Flow and Immersion, the results of which show higher scores when diegetic sound is present than when it is not. *Prima facie*, this would indicate that diegetic sound does indeed have an immersive effect in the case of FPS games. Music also appears to increase immersion, while reducing tension and negative affect, at the expense of a reduction in the experience of flow within gameplay.

Secondly, the psychophysiological data do not support the subjective results, but are instead both inconclusive and lacking statistical significance (see Table 2). If we maintain the assumption that physiological evidence, in these circumstances, can be used to confirm the subjective evidence, then there are several potential explanations for the lack of correlation between the two result sets. Further analysis and experimentation will be required to explain this disparity. Some initial possible explanations (assuming a valid experiment design and implementation) include:

1. *The GEQ incorporates distortions derived from the retrospective "storytelling" context of the questionnaire.*

2. *The physiological data, gathered over 10 minutes of play, contains too much noise to produce a significant result.* It must be noted that the data analyzed here was accumulated over one game session and even after inspection of histograms and logarithmic correction not all measurements were perfectly normally distributed. Even though a non-parametric statistical analysis or a range correction of physiological responses could be conducted, it is unlikely that this will show significant effects over the 10 minute timescale used. Connecting physiological response data to game events using more precise phasic measurements, as described in Nacke et al., [30] could yield more insight into the emotional effects of sound. This level of detail can be achieved but it would need an additional method for recording subjective

[3] N = 29 (after data reduction).

responses at the same event level precision to be correlated with.

3. *The subjectively reported experience is a function of the modulation of emotions within a smaller time scale than that used in the analysis of psychophysiological data*. This means that the emotional net effect may be the same, but the details of emotional dynamics produce different subjective experiences as reported by the GEQ. As analogy: a flat sea and a sea with big waves may have the same mean level, but one makes for much better surfing than the other. This might be detectable by derived measures form the current data set.

4. *The subjectively reported results are not measurable using our equipment and methods.* In particular, the source of the GEQ components reported in Table 1 may have a different psychological explanation than that captured by the arousal/valence model of emotion. This consideration raises the need for more thorough ongoing conceptual investigations of terms such as immersion, presence, flow, challenge and fun (as started in [31]). Based upon a richer range of linguistic and conceptual distinctions, it may be possible to devise experiments having more discriminating power among the range of descriptive models thus created. In particular, these are complex concepts used in different ways by different authors, and it may not be the case that they have simple mappings to instantaneous emotions measured by psychophysiological techniques. Explanatory theories then need to move to higher levels in modeling the structuring of a series of measurable emotions, related to perceptions and player actions, to provide a more systemic account of the foundations of the quality of play experience, as suggested by Lindley and Sennersten. [32]

These questions must be addressed by ongoing research. To our surprise, our research contradicts the results presented by Shilling et al. [18], who indicated a strong correlation between sounds and physiologically elicited emotions. Unfortunately, Shilling et al. did not report direct values of their measures that would allow a direct comparison. It remains for more thorough future analysis to find greater scientific evidence for a relationship between sound and psychophysiological measures. Our future aim is to investigate this within our research.

## Acknowledgements

The research reported in this paper has been partially funded by the FUGA (FUn of GAming) EU FP6 research project (NEST-PATH-028765). We thank our FUGA colleagues, especially Niklas Ravaja, Matias Kivikangas, and Simo Järvelä, for many stimulating inputs to this work. We would also like to thank laboratory assistant Dennis Sasse for helping in the execution of the experiment.

---

## References

[1] id Software, *Doom* series, Activision, (1993-2004).
[2] id Software, *Quake* series, Activision, (1996-2005).
[3] Valve Software, *Half-Life* series, Electronic Arts, (1998-2004).
[4] Grimshaw, M., *The Acoustic ecology of the first-person shooter: The player, sound and immersion in the first-person shooter computer game*, Saarbrücken, VDM Verlag Dr. Mueller e.K., (2008).
[5] McMahan, A., *Immersion, engagement, and presence: A new method for analyzing 3-D video games*, in M. J. P. Wolf & B. Perron (Eds.), The Video Game Theory Reader, New York, Routledge, 67–87 (2003).
[6] Kearney, P. R., & Pivec, M., *Immersed and how? That is the question*, Game in' Action, Sweden, (2007, June 13—15).
[7] Pine, B. J., & Gilmore, J. H., *The experience economy: Work is theatre & every business a stage*, Boston, Harvard Business School Press, (1999).
[8] Ermi, L., & Mäyrä, F., *Fundamental components of the gameplay experience: Analysing immersion*, Changing Views – Worlds in Play, Toronto, (June 16—20, 2005).
[9] Murray, J. H., *Hamlet on the holodeck: The future of narrative in cyberspace*, Cambridge, The MIT Press, (2000).
[10] Garcia, J. M., *From heartland values to killing prostitutes: An overview of sound in the video game Grand Theft Auto Liberty City Stories*, Audio Mostly 2006, Piteå, Sweden, (October 11—12, 2006).
[11] Carr, D., *Space, navigation and affect*, in Computer Games: Text, Narrative and Play, Cambridge, Polity, 59-71 (2006).
[12] Laurel, B., *Computers as theatre*, New York, Addison-Wesley, (1993).
[13] Jørgensen, K., *On the functional aspects of computer game audio*, Audio Mostly 2006, Piteå, Sweden, (October 11—12, 2006).
[14] Murphy, D., & Pitt, I., *Spatial sound enhancing virtual story telling,* Lecture Notes in Computer Science, 2197*,* 20–29 (2001).
[15] Grimshaw, M. & Schott, G., *A Conceptual framework for the design and analysis of first-person shooter audio and its potential use for game engines*, International Journal of Computer Games Technology, 2008, (2008).
[16] Stockburger, A., *The rendered arena: Modalities of space in video and computer games*, unpublished PhD thesis, London, University of the Arts, (2006).
[17] Grimshaw, M., *The Resonating spaces of first-person shooter games*, Proceedings of The 5th International Conference on Game Design and Technology, Liverpool, (November 14—15, 2007).
[18] Shilling, R., Zyda, M., & Wardynski, E. C., *Introducing emotion into military simulation and videogame design: America's Army: Operations and VIRTE*, GameOn, London, (2002).
[19] MOVES Institute, *America's army*, Monterey, Naval Postgraduate School, (2002).
[20] Cacioppo, J.T., Tassinary, L.G. and Berntson, G.G. *Psychophysiological science*, Handbook of psychophysiology, 3-26 (2007).
[21] Cacioppo, J.T., *Handbook of Psychophysiology*, Cambridge University Press, (2007).
[22] Lang, P. J., *The emotion probe. Studies of motivation and attention*, American Psychologist, 50, 372–385 (1995).
[23] Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O., *Looking at pictures: Affective, facial, visceral, and behavioral reactions*, Psychophysiology, 30, 261–273 (1993).
[24] Valve Software, *Half-Life 2*, Electronic Arts, (2004).
[25] Schmitz, C., *LimeSurvey* (v1.70) from http://www.limesurvey.org (2008).
[26] Fridlund, & Cacioppo., *Guidelines for human electromyographic research,* Psychophysiology, 23(5), 567-589 (1986).
[27] IJsselsteijn, W.A., Poels, K. and de Kort, Y.A.W., *Measuring player experiences in digital games. Development of*

*the Game Experience Questionnaire (GEQ)*, Manuscript in preparation.
[28] Nacke, L. and Lindley, C., *Boredom, Immersion, Flow - A Pilot Study Investigating Player Experience,* IADIS Gaming 2008: Design for Engaging Experience and Social Interaction, IADIS, Amsterdam, The Netherlands, (2008).
[29] Ravaja, N., Turpeinen, M., Saari, T., Puttonen, S. and Keltikangas-Jarvinen, L., *The Psychophysiology of James Bond: Phasic Emotional Responses to Violent Video Game Events*, Emotion, 8 (1), 114-120 (2008).
[30] Nacke, L., Lindley, C. and Stellmach, S., *Log who's playing: psychophysiological game analysis made easy through event logging*, International conference on Fun and Games 2008, Springer, Eindhoven, The Netherlands, (2008).
[31] Lindley, C.A., Nacke, L. and Sennersten, C., *What does it mean to understand gameplay?* First Symposium on Ludic Engagement Designs for All, Aalborg University, Esbjerg, Denmark, (2007).
[32] Lindley C. A. and Sennersten C., 2006 *A Cognitive Framework for the Analysis of Game Play: Tasks, Schemas and Attention Theory*, Workshop on the Cognitive Science of Games and Game Play, The 28th Annual Conference of the Cognitive Science Society, Vancouver, Canada, (July 26—29, 2006).